

CHRONIC DISEASE COST CALCULATOR

TECHNICAL APPENDIX

Version 1

October 2008

Centers for Disease Control and Prevention (CDC)



Developed by RTI International*



In collaboration with the Agency for Healthcare Research and Quality (AHRQ),
the National Association of Chronic Disease Directors (NACDD),
and the National Pharmaceutical Council (NPC)



* RTI International is a trade name of Research Triangle Institute.

Table of Contents

Default Source Data Sets	3
<i>Medicaid Enrollment</i>	3
<i>Prevalence and Per-Person Costs</i>	3
Methodology	4
<i>Medicaid Enrollment</i>	4
<i>Prevalence</i>	5
<i>Per-Person Costs</i>	6
<i>Calculated Costs</i>	8
External Review	8
Glossary of Terms	9
References	10
Appendix: ICD-9 Codes used in the Analysis	11

Default Source Data Sets

Medicaid Enrollment

Medicaid Statistical Information System (MSIS)

Medicaid enrollment data came from the Medicaid Statistical Information System (MSIS) State Summary Fiscal Year 2004.¹ MSIS data are used by the Centers for Medicare & Medicaid Services (CMS) to produce Medicaid program characteristics and utilization information for the states. The purpose of MSIS is to collect, manage, analyze and disseminate information on eligibles, beneficiaries, utilization and payment for services covered by State Medicaid programs. Fiscal Year 2004 was the most current data available at the time of calculator development.

Prevalence and Per-Person Costs

Medicaid Analytic eXtract (MAX)

Estimates for the institutionalized population, which are not available in other data sources, were derived from the Medicaid Analytic eXtract (MAX). MAX is a uniform dataset containing Medicaid eligibility, utilization, and payment information that is created by CMS from MSIS data submitted by the states.² All states have been included in the MAX files since 1999. MAX files include claims-level data for all services covered by Medicaid, organized as inpatient, long-term care (LTC), other provider, and prescription drug files. The MAX data also contain a person-summary file, which includes monthly Medicaid eligibility and managed care enrollment data, limited demographic information, and aggregate utilization and expenditure data. MAX data were provided by CMS under a Memorandum of Understanding with RTI International.

Four states were analyzed: Illinois, Indiana, Kansas, and Louisiana. Criteria for selecting states for the Medicaid claims analyses included: (1) having a low percentage of Medicaid beneficiaries enrolled in capitated managed care plans, (2) having reasonably complete diagnosis code reporting in Medicaid claims data, and (3) variation in mortality due to heart disease and stroke. Cost constraints limited the number of states. The study population included 100% of beneficiaries with at least 1 month of enrollment in fee-for-service Medicaid during 2001. Beneficiaries who were only enrolled in capitated Medicaid managed care during the year were excluded. Eligibility and utilization data for periods of enrollment in capitated managed care were excluded from the analyses for those beneficiaries with periods of enrollment in both fee-for-service Medicaid and capitated managed care during the year.

Diseases were defined using *International Classification of Diseases, Ninth Revision* (ICD-9) diagnosis codes in claims data (see Appendix). Hypertension and diabetes were also considered present if a person had a prescription drug claim for drugs in those therapeutic classes. Individuals with at least one inpatient, long term care, or prescription drug claim that met the criteria for a condition were classified as having that condition. For all other types of service, the individual was required to have at least two claims on different dates that met the condition criteria. A full list of ICD-9 and prescription drug codes is available upon request. All

expenditure data were inflated to 2007 dollars using the Medical Care Consumer Price Index to reflect more current dollar values.

Medical Expenditure Panel Survey (MEPS)

Data were pooled from the 2001 through 2005 Medical Expenditure Panel Survey (MEPS) Consolidated Data Files, a nationally representative survey of the civilian non-institutionalized population that provides data on annual medical expenditures, sources of payment and insurance coverage for each participant.³ Diseases were defined using ICD-9 codes based on self-reported diseases that were transcribed by professional coders and reported in the MEPS Medical Conditions files for years 2001 through 2005 (see Appendix).

The combined four-year MEPS sample included 165,892 persons of all ages living in the U.S. The study population consisted of anyone covered by Medicaid at any point during the interview year (N=32,001). Estimates for both prevalence and costs have been adjusted to be nationally representative using MEPS sampling weights for years 2001 through 2005. All expenditure data were inflated to 2007 dollars using the Medical Care Consumer Price Index to reflect more current dollar values.

Methodology

This section describes the methods used to generate the default values in the calculator.

The prevalence and per-person costs are built up from weighted averages of estimates by sex by age (e.g., males age 65+). The weights are derived from the Medicaid population distribution in each state. This assures that the subgroup and total population prevalence rates and costs will always be consistent with one another for any Medicaid population the user enters (i.e., the sum of Medicaid disease costs by sex and the sum of Medicaid disease costs by age both equal the total Medicaid disease costs). However, small sample sizes in MEPS preclude reporting of prevalence and costs by sex by age subgroups. Where sample sizes permit, we report prevalence and costs separately by sex and by adult age group.

Medicaid Enrollment

Medicaid enrollment was taken from the MSIS State Summary Fiscal Year 2004. The age categories reported by MSIS were slightly different than the age categories used in the analysis. In order to calculate the number of beneficiaries 18-44 years of age, we added one quarter of the total number of beneficiaries age 15-18 to the total number of beneficiaries age 19-44.

The total number of beneficiaries reported by MSIS included those with unknown age or sex. The number of beneficiaries with unknown age were split into age categories 0-17, 18-44, 45-64, and 65 and over based on the share of the beneficiary population in each age group for those with observed age at the state level. The same method was applied to those with missing sex.

MSIS does not report Medicaid enrollment by sex by age. We estimated the share of males and females within each of our age categories in the MAX data pooled across states. We then applied this sex distribution by age category to each of the state enrollment populations in

MSIS. The resulting Medicaid population distributions by sex by age were used as weights to calculate weighted average prevalence and per-person costs by sex, by age, and for the overall Medicaid population in each state.

Beneficiaries who were only enrolled in capitated Medicaid managed care during the year were not excluded from the enrollment totals.

Prevalence

Prevalence was defined as the percentage of the Medicaid population reporting treatment for or problems with the disease in the previous (i.e., interview) year.

- § The MEPS 2001-2005 Condition files were used to define prevalence.
- § Prevalence in the MEPS data was based on self-reports of being treated for or being bothered by the disease within the interview year.
 - “Conditions can be added to the MEPS condition roster in one of several ways. Most directly, the condition can be identified as the reason reported by the household respondent for a particular medical event (hospital stay, outpatient visit, emergency room visit, home health episode, prescribed medication purchase, or medical provider visit). Second, the condition may have been reported as the reason for one or more episodes of disability days. Finally, the condition may have been reported by the household level respondent as a condition “bothering” the person during the reference period.”⁴
 - For example, a person in remission for cancer but who visited a doctor for a follow-up appointment during the interview year would be included in the cancer prevalence rate. A person with a history of hypertension but who did not have any medical treatment or report being bothered by his/her hypertension during the interview year would not be included in the hypertension prevalence rate.
 - MEPS prevalence was based on a shorter time frame (one year) than other self-reported prevalence estimates like the Behavioral Risk Factor Surveillance System (BRFSS), which asks if the person has ever been diagnosed with the disease.
 - *Persons identified as having the disease in the MEPS have likely had recent treatment or problems with the disease and thus were likely more resource-intensive than those included in alternative prevalence definitions based on a history of the disease who have not sought treatment recently.*
- § All default prevalence rates were weighted using the MEPS sampling weights to be nationally representative.
- § Prevalence rates were calculated by sex by age subgroups at the Census region level, the most detailed geographic data available in the public use MEPS. We only report weighted average prevalence separately by sex and by age group when at least 100 unweighted observations with the disease are available in MEPS for that subgroup and region. Where sample sizes are too small, weighted average prevalence from the national level separately by sex and by age was used. In cases where sample sizes were too small even

at the national level for estimates by sex or age (e.g., CHF), the calculator does not report output by these subgroups.

- § Prevalence in the MEPS data did not include those in LTC. Therefore, all MEPS prevalence rates were scaled to account for missing LTC prevalence. The scaling factor, drawn from the MAX data pooled across the four states, was the ratio of overall prevalence including the LTC population to prevalence in the non-LTC population. Separate ratios were used for each sex/age subpopulation.
- § Children (i.e., ages 0 to 17) were assumed to have zero prevalence for these diseases. The small number of children with the disease in MEPS prevented reliable prevalence estimates for this age group. Actual prevalence rates for children in the pooled MAX data were very small: heart disease = 0.25%, CHF = 0.03%, hypertension = 1.5%, stroke = 0.04%, diabetes = 0.20%, and cancer = 0.12%.

Per-Person Costs

Econometric modeling was used to estimate the costs associated with selected chronic diseases. The econometric approach allowed the model to predict expenditures with and without the chronic disease and isolated the impact of disease while controlling for the presence of other important drivers of medical expenditures that are not influenced by health policy such as age. Models were run for the total Medicaid population. Sample sizes are too small to estimate separate models by sex by age. Results reported by sex and by age category were generated from weighted averages of simulations by sex by age using the regression model for the entire Medicaid population (see below for description of simulations). Nonlinearities in the overall model will generate different average costs per person for the different sex/age groups.

- § A two-part regression model was used to predict annual per-person costs. The first part of the two-part model used logistic regression to predict the probability of any expenditures. The second part of the two-part model analyzed expenditures conditional on having any expenditures (i.e., any utilization). To select the appropriate cost estimation model for the second part of the two-part model, we used the algorithm for choosing among alternative non-linear estimators recommended by Manning and Mullahy.⁵ The results of this algorithm indicated that we should use a generalized linear model (GLM) with a gamma distribution and a log link in the second stage to analyze non-zero expenditures.
- § The dependent variable in each regression was restricted to the annual medical expenditures paid by Medicaid.
- § All regressions included the following variables: indicators for heart disease, CHF, hypertension, stroke, diabetes, and cancer; age; age squared; sex; race/ethnicity; education; family income; and indicators for other diseases including injuries, dyslipidemia, HIV/AIDS, pneumonia, asthma, chronic obstructive pulmonary disease, depression, other mental health/substance abuse, arthritis, back problems, skin disorders, renal failure, and pregnancies.
- § Hypertension and diabetes are risk factors for other diseases included in the cost estimation model; therefore, the full model likely underestimated the true costs of hypertension and diabetes. To capture the costs of diseases on the causal pathway, the estimates for hypertension and diabetes are from alternate models that omitted controlling for diseases linked to hypertension and diabetes.

- The hypertension model dropped heart disease, stroke, CHF, and renal failure.
 - The diabetes model dropped heart disease, stroke, CHF, hypertension, renal failure, and dyslipidemia.
- § Separate two-part models were used for each of the following types of service: inpatient, ambulatory (outpatient, office-based, emergency, and other), and prescription. The per-patient costs were then summed across services to get the total.
- § Relative standard errors for the national MEPS disease expenditures range from 21% to 35% across diseases.
- § Confidential MEPS data that identified the largest 30 states and 9 census divisions were utilized to generate state-specific per-person cost estimates.
- Medicaid sample sizes were not large enough by state to replicate the full analysis for each state.
 - We regressed log (positive) medical expenditures on the variables in the model plus state/census division dummies.
 - The coefficients on the dummies provided measures of the differences in average medical costs across states that were used to scale the national estimates to make them state-specific.
- § MEPS expenditures did not include LTC. The following steps were taken to include LTC costs in the per-person cost estimates:
- Ordinary least squares (OLS) regression was run on the MAX data pooled across the four states with LTC expenditures as the dependent variable. Many alternative specifications were tested for the MAX analysis, including lognormal and GLM gamma models to account for skewness in the distribution. The chosen specification produced the most stable and, in most cases, most conservative estimates. Use of OLS regression is also justified because of the large number of observations in the MAX files.
 - The regressions included the following variables: indicators for heart disease, CHF, hypertension, stroke, diabetes, and cancer; age; age squared; sex; race/ethnicity; and indicators for other diseases including injuries, dyslipidemia, HIV/AIDS, pneumonia, asthma, chronic obstructive pulmonary disease, depression, other mental health/substance abuse, arthritis, back problems, skin disorders, renal failure, pregnancies, organic psychoses (excluding alcohol and drug psychoses), mental retardation, and hereditary and degenerative diseases of the central nervous system.
 - This provided average LTC costs per person with the disease.
 - Average LTC costs per person with the disease were converted to nationally-representative estimates using a weighted average of the state multipliers for the four MAX states where the weights were the share of observations in the pooled data from each state.
 - The national LTC per-person cost estimate was added to the MEPS estimates before converting to state-level estimates using the state/census division multipliers.
- § Costs per person attributable to each disease were calculated using the following method that minimizes double-counting of expenditures across diseases.⁶

- First, every unique combination of the chronic diseases observed in the data was identified (e.g., stroke, stroke with hypertension, and stroke with CHF are three of the unique combinations).
- Second, expenditures were predicted for each individual.
- Then, for each unique combination of diseases, we subtracted from step two the predicted expenditures for an otherwise identical person without the combination of diseases. This provides an estimate of the costs attributable to every unique combination of diseases.
- Next, the coefficients of the diseases from the second part of the two part econometric model were used as importance weights to redistribute costs associated with jointly-occurring diseases to constituent diseases (e.g., to redistribute the costs of stroke with hypertension back to stroke and hypertension separately).
- Finally, the application averages the redistributed costs over the population with each disease.
- There should not be any double-counting of costs between the LTC and non-LTC analyses because the former only includes the LTC component of costs, which are excluded entirely from the non-LTC analyses.

Calculated Costs

- § The total number of Medicaid beneficiaries with a disease is calculated by multiplying the number of beneficiaries by the prevalence rate.
- § The total annual Medicaid costs attributable to a disease are calculated by multiplying the average annual costs per person with the disease by the number of beneficiaries with the disease.
- § The total Medicaid costs for all selected diseases are calculated by summing the total annual Medicaid costs attributable to each disease. However, as described above, the cost estimates for hypertension and diabetes include the costs of complications such as heart disease, CHF and stroke. *The sum of costs over selected diseases that include hypertension and diabetes could overestimate the costs associated with all the selected diseases.*

External Review

The methods used in generating estimates for the calculator were reviewed by an external panel of subject matter experts.

Glossary of Terms

Attributable: The portion of expenditures directly associated with a disease.

BRFSS: Behavioral Risk Factor Surveillance System

CHF: Congestive heart failure

CMS: Centers for Medicare & Medicaid Services

Generalized linear model: A type of non-linear estimation used to compute the relationship between a set of independent variables and a dependent variable that has a non-normal distribution. In this case, the independent variables are the demographics and the disease indicators and the dependent variable is strictly positive Medicaid expenditures, which is positively skewed.

ICD-9: International Classification of Diseases, Ninth Revision

Logistic regression: A type of non-linear estimation used to compute the relationship between a set of independent variables and a discrete dependent variable. In this case, the independent variables are the demographics and the disease indicators and the dependent variable is the probability of having positive Medicaid expenditures.

LTC: Long-term care

MAX: Medicaid Analytic eXtract

MEPS: Medical Expenditure Panel Survey

MSIS: Medicaid Statistical Information System

NHANES: National Health and Nutrition Examination Survey

Per-person costs: Average Medicaid costs per person with the disease attributable to the disease. These numbers represent the extent to which Medicaid expenditures per person with the disease would be lower in the absence of the disease, all else constant.

Prevalence: The percentage of the Medicaid population reporting treatment for or problems with the disease in the previous (i.e., interview) year.

References

1. Centers for Medicare & Medicaid Services.
http://www.cms.hhs.gov/MedicaidDataSourcesGenInfo/02_MSISData.asp. Last accessed May 6, 2008.
2. Centers for Medicare & Medicaid Services.
http://www.cms.hhs.gov/MedicaidDataSourcesGenInfo/07_MAXGeneralInformation.asp
Accessed November 16, 2006.
3. Full documentation and data is available from Agency for Healthcare Research and Quality,
<http://www.meps.ahrq.gov/>. Accessed November 16, 2006.
4. Agency for Healthcare Research and Quality.
http://www.meps.ahrq.gov/mepsweb/data_stats/download_data/pufs/h96/h96doc.pdf,
page C-8. Last accessed October 22, 2008.
5. Manning WG, Mullahy J. Estimating log models: To transform or not to transform? *Journal of Health Economics* 2001;20:461-494.
6. Trogdon JG, Finkelstein EA, Hoerger TJ. Use of econometric models to estimate expenditure shares. *Health Services Research* 2008;43(4):1442-1451.

Appendix: ICD-9 Codes used in the Analysis

Appendix Table 1. ICD-9 Codes Used to Define Diseases in MEPS

Heart Disease	
Heart valve disorders	V422 V433 3940 3941 3942 3949 3950 3951 3952 3959 3960 3961 3962 3963 3968 3969 3970 3971 3979 4240 4241 4242 4243 42490 42491 42499 7852 7853
Peri-, endo-, and myocarditis, cardiomyopathy	3910 3911 3912 3918 3919 3920 393 3980 39890 39899 4200 42090 42091 42099 4210 4211 4219 4220 42290 42291 42292 42293 42299 4230 4231 4232 4238 4239 4250 4251 4252 4253 4254 4255 4257 4258 4259 4290
Acute myocardial infarction	4100 41000 41001 41002 4101 41010 41011 41012 4102 41020 41021 41022 4103 41030 41031 41032 4104 41040 41041 41042 4105 41050 41051 41052 4106 41060 41061 41062 4107 41070 41071 41072 4108 41080 41081 41082 4109 41090 41091 41092
Coronary atherosclerosis and other heart disease	V4581 V4582 4110 4111 4118 41181 41189 412 4130 4131 4139 4140 41400 41401 4148 4149
Pulmonary heart disease	4150 4151 41519 4160 4161 4168 4169 4170 4171 4178 4179
Other and ill-defined heart disease	41410 41411 41412 41419 4291 4292 4293 4295 4296 42971 42979 42981 42982 42989 4299
Conduction disorders	V450 V4500 V4501 V4502 V4509 V533 V5331 V5332 V5339 4260 42610 42611 42612 42613 4262 4263 4264 42650 42651 42652 42653 42654 4266 4267 42681 42689 4269 4270 4271 4272 42731 42732 42760 42761 42769 42781 42789 4279 7850 7851
Cardiac dysrhythmias	42741 42742 4275
Cardiac arrest and ventricular fibrillation	402 404
Hypertension with heart disease	
Congestive Heart Failure	
Congestive heart failure, nonhypertensive	39891 4280 4281 42820 42821 42822 42823 42830 42831 42832 42833 42840 42841 42842 42843 4289
Hypertension	
Hypertension without heart disease	401 403 405
Stroke	
Acute cerebrovascular disease	430 431 4320 4321 4329 43301 43311 43321 43331 43381 43391 4340 43400 43401 4341 43410 43411 4349 43490 43491 436
Occlusion or stenosis of precerebral arteries	4330 43300 4331 43310 4332 43320 4333 43330 4338 43380 4339 43390
Other and ill-defined cerebrovascular disease	4370 4371 4373 4374 4375 4376 4377 4378 4379

Appendix Table 1. (cont.)	
Late effects of cerebrovascular disease	438 4380 43810 43811 43812 43819 43820 43821 43822 43830 43831 43832 43840 43841 43842 43850 43851 43852 43853 4386 4387 43881 43882 43883 43884 43885 43889 4389
Diabetes	
Diabetes mellitus with and without complications	250
Cancer	
Cancer of colon	V1005 1530 1531 1532 1533 1534 1535 1536 1537 1538 1539 1590 2303
Cancer of rectum and anus	V1006 1540 1541 1542 1543 1548 2304 2305 2306
Cancer of bronchus, lung	V1011 1622 1623 1624 1625 1628 1629 2312
Cancer of breast	V103 1740 1741 1742 1743 1744 1745 1746 1748 1749 1750 1759 2330
Cancer of prostate	V1046 185 2334
Melanomas of skin	V1082 1720 1721 1722 1723 1724 1725 1726 1727 1728 1729
Other non-epithelial cancer of skin	V1083 1730 1731 1732 1733 1734 1735 1736 1737 1738 1739 2320 2321 2322 2323 2324 2325 2326 2327 2328 2329
Cancer of Cervix	V1041 1800 1801 1808 1809 2331 7950
Cancer of head and neck	V1001 V1002 V1021 1400 1401 1403 1404 1405 1406 1408 1409 1410 1411 1412 1413 1414 1415 1416 1418 1419 1420 1421 1422 1428 1429 1430 1431 1438 1439 1440 1441 1448 1449 1450 1451 1452 1453 1454 1455 1456 1458 1459 1460 1461 1462 1463 1464 1465 1466 1467 1468 1469 1470 1471 1472 1473 1478 1479 1480 1481 1482 1483 1488 1489 1490 1491 1498 1499 1600 1601 1602 1603 1604 1605 1608 1609 1610 1611 1612 1613 1618 1619 1950 2300 2310
Cancer of esophagus	V1003 1500 1501 1502 1503 1504 1505 1508 1509 2301
Cancer of stomach	V1004 1510 1511 1512 1513 1514 1515 1516 1518 1519 2302
Cancer of liver and intrahepatic bile duct	V1007 1550 1551 1552 2308
Cancer of pancreas	1570 1571 1572 1573 1574 1578 1579
Cancer of other GI organs, peritoneum	V1000 V1009 1520 1521 1522 1523 1528 1529 1560 1561 1562 1568 1569 1580 1588 1589 1591 1598 1599 2307 2309
Cancer of other respiratory and intrathoracic	V1012 V1020 V1022 1620 1630 1631 1638 1639 1650 1658 1659 2311 2318 2319
Cancer of bone and connective tissue	1700 1701 1702 1703 1704 1705 1706 1707 1708 1709 1710 1712 1713 1714 1715 1716 1717 1718 1719
Cancer of ovary	V1043 1830
Cancer of other female genital organs	V1040 V1044 181 1832 1833 1834 1835 1838 1839 1840 1841 1842 1843 1844 1848 1849 2333

Appendix Table 1. (cont.)

Cancer of testis	V1047 1860 1869
Cancer of other male genital organs	V1045 V1048 V1049 1871 1872 1873 1874 1875 1876 1877 1878 1879 2335 2336
Cancer of bladder	V1051 1880 1881 1882 1883 1884 1885 1886 1887 1888 1889 2337
Cancer of kidney and renal pelvis	V1052 1890 1891
Cancer of other urinary organs	V1050 V1059 1892 1893 1894 1898 1899 2339
Cancer of brain and nervous system	V1085 V1086 1910 1911 1912 1913 1914 1915 1916 1917 1918 1919 1920 1921 1922 1923 1928 1929
Cancer of thyroid	V1087 193
Hodgkin's disease	V1072 20100 20101 20102 20103 20104 20105 20106 20107 20108 20110 20111 20112 20113 20114 20115 20116 20117 20118 20120 20121 20122 20123 20124 20125 20126 20127 20128 20140 20141 20142 20143 20144 20145 20146 20147 20148 20150 20151 20152 20153 20154 20155 20156 20157 20158 20160 20161 20162 20163 20164 20165 20166 20167 20168 20170 20171 20172 20173 20174 20175 20176 20177 20178 20190 20191 20192 20193 20194 20195 20196 20197 20198
Non-Hodgkin's lymphoma	V1071 V1079 20000 20001 20002 20003 20004 20005 20006 20007 20008 20010 20011 20012 20013 20014 20015 20016 20017 20018 20020 20021 20022 20023 20024 20025 20026 20027 20028 20080 20081 20082 20083 20084 20085 20086 20087 20088 20200 20201 20202 20203 20204 20205 20206 20207 20208 20210 20211 20212 20213 20214 20215 20216 20217 20218 20220 20221 20222 20223 20224 20225 20226 20227 20228 20280 20281 20282 20283 20284 20285 20286 20287 20288 20290 20291 20292 20293 20294 20295 20296 20297 20298
Leukemias	V1060 V1061 V1062 V1063 V1069 20240 20241 20242 20243 20244 20245 20246 20247 20248 2031 20310 20311 2040 20400 20401 2041 20410 20411 2042 20420 20421 2048 20480 20481 2049 20490 20491 2050 20500 20501 2051 20510 20511 2052 20520 20521 2053 20530 20531 2058 20580 20581 2059 20590 20591 2060 20600 20601 2061 20610 20611 2062 20620 20621 2068 20680 20681 2069 20690 20691 2070 20700 20701 2071 20710 20711 2072 20720 20721 2078 20780 20781 2080 20800 20801 2081 20810 20811 2082 20820 20821 2088 20880 20881 2089 20890 20891
Multiple myeloma	2030 20300 20301 2038 20380 20381

Appendix Table 1. (cont.)

Cancer, other and unspecified primary

V1029 V1081 V1084 V1088 V1089 V109 V711
1640 1641 1642 1643 1648 1649 1760 1761
1762 1763 1764 1765 1768 1769 1900 1901
1902 1903 1904 1905 1906 1907 1908 1909
1940 1941 1943 1944 1945 1946 1948 1949
1951 1952 1953 1954 1955 1958 20230 20231
20232 20233 20234 20235 20236 20237 20238
20250 20251 20252 20253 20254 20255 20256
20257 20258 20260 20261 20262 20263 20264
20265 20266 20267 20268 2340 2348 2349
7951

Secondary malignancies

1960 1961 1962 1963 1965 1966 1968 1969
1970 1971 1972 1973 1974 1975 1976 1977
1978 1980 1981 1982 1983 1984 1985 1986
1987 19881 19882 19889

Malignant neoplasm without specification of site

1990 1991

Maintenance chemotherapy, radiotherapy

V580 V581 V661 V662 V671 V672 V67

Appendix Table 2. ICD-9 Codes Used to Define Diseases in MAX Data

Heart Disease	391	392	393	394	395	396
		397	398	402	404	410
		411	412	413	414	415
		416	417	418	419	420
		421	422	423	424	425
		426	427	429	V422	V433
		V450	V4500	V4501	V4502	V4509
		V4581	V4582	V533	V5331	V5332
		V5339				
Congestive Heart Failure	428					
Hypertension*	401	403	405			
Stroke	430	431	432	433	434	436
		437	438			
Diabetes*	25000	25001	25002	25003	25010	25011
		25012	25013	25020	25021	25022
		25023	25030	25031	25032	25033
		25040	25041	25042	25043	25050
		25051	25052	25053	25060	25061
		25062	25063	25070	25071	25072
		25073	25080	25081	25082	25083
		25090	25091	25092	25093	3572
		36201	36202	36641	3260	
Cancer						
Colorectal cancer	1530	1531	1532	1533	1534	1535
		1536	1537	1538	1539	1540
		1541	1542	1543	1548	2303
		2304	2305	2306		
Lung cancer	1622	1623	1624	1625	1628	1629
		2312				
Breast cancer	1740	1741	1742	1743	1744	1745
		1746	1748	1749	1750	1759
		2330				
Prostate cancer	185	2334				
Skin cancer	1720	1721	1722	1723	1724	1725
		1726	1727	1728	1729	1730
		1731	1732	1733	1734	1735
		1736	1737	1738	1739	2320
		2321	2322	2323	2324	2325
		2326	2327	2328	2329	
Ovarian cancer	1830	1832	1833	1834	1835	1838
		1839	V1043			
Cervical cancer	1800	1801	1808	1809	2331	
Other cancers	1400	1401	1403	1404	1405	1406
		1408	1409	1410	1411	1412
		1413	1414	1415	1416	1418
		1419	1420	1421	1422	1428
		1429	1430	1431	1438	1439
		1440	1441	1448	1449	1450
		1451	1452	1453	1454	1455
		1456	1458	1459	1460	1461
		1462	1463	1464	1465	1466
		1467	1468	1469	1470	1471

Appendix Table 2. (cont.)

1472	1473	1478	1479	1480
1481	1482	1483	1488	1489
1490	1491	1498	1499	1500
1501	1502	1503	1504	1505
1508	1509	1510	1511	1512
1513	1514	1515	1516	1518
1519	1520	1521	1522	1523
1528	1529	1550	1551	1552
1560	1561	1562	1568	1569
1570	1571	1572	1573	1574
1578	1579	1580	1588	1589
1590	1591	1598	1599	1600
1601	1602	1603	1604	1605
1608	1609	1610	1611	1612
1613	1618	1619	1620	1630
1631	1638	1639	1640	1641
1642	1643	1648	1649	1650
1658	1659	1700	1701	1702
1703	1704	1705	1706	1707
1708	1709	1710	1712	1713
1714	1715	1716	1717	1718
1719	1760	1761	1762	1763
1764	1765	1768	1769	179
181	1820	1821	1828	1840
1841	1842	1843	1844	1848
1849	1860	1869	1871	1872
1873	1874	1875	1876	1877
1878	1879	1880	1881	1882
1883	1884	1885	1886	1887
1888	1889	1890	1891	1892
1893	1894	1898	1899	1900
1901	1902	1903	1904	1905
1906	1907	1908	1909	1910
1911	1912	1913	1914	1915
1916	1917	1918	1919	1920
1921	1922	1923	1928	1929
193	1940	1941	1943	1944
1945	1946	1948	1949	1950
1951	1952	1953	1954	1955
1958	1960	1961	1962	1963

Appendix Table 2. (cont.)

1965	1966	1968	1969	1970
1971	1972	1973	1974	1975
1976	1977	1978	1980	1981
1982	1983	1984	1985	1986
1987	19881	19882	19889	1990
1991	20000	20001	20002	20003
20004	20005	20006	20007	20008
20010	20011	20012	20013	20014
20015	20016	20017	20018	20020
20021	20022	20023	20024	20025
20026	20027	20028	20080	20081
20082	20083	20084	20085	20086
20087	20088	20100	20101	20102
20103	20104	20105	20106	20107
20108	20110	20111	20112	20113
20114	20115	20116	20117	20118
20120	20121	20122	20123	20124
20125	20126	20127	20128	20140
20141	20142	20143	20144	20145
20146	20147	20148	20150	20151
20152	20153	20154	20155	20156
20157	20158	20160	20161	20162
20163	20164	20165	20166	20167
20168	20170	20171	20172	20173
20174	20175	20176	20177	20178
20190	20191	20192	20193	20194
20195	20196	20197	20198	20200
20201	20202	20203	20204	20205
20206	20207	20208	20210	20211
20212	20213	20214	20215	20216
20217	20218	20220	20221	20222
20223	20224	20225	20226	20227
20228	20230	20231	20232	20233
20234	20235	20236	20237	20238
20240	20241	20242	20243	20244
20245	20246	20247	20248	20250
20251	20252	20253	20254	20255
20256	20257	20258	20260	20261
20262	20263	20264	20265	20266
20267	20268	20280	20281	20282
20283	20284	20285	20286	20287
20288	20290	20291	20292	20293
20294	20295	20296	20297	20298
20300	20301	20310	20311	20380
20381	20400	20401	20410	20411
20420	20421	20480	20481	20490
20491	20500	20501	20510	20511
20520	20521	20530	20531	20580
20581	20590	20591	20600	20601
20610	20611	20620	20621	20680

Appendix Table 2. (cont.)

20681	20690	20691	20700	20701
20710	20711	20720	20721	20780
20781	20800	20801	20810	20811
20820	20821	20880	20881	20890
20891	2300	2301	2302	2307
2308	2309	2310	2311	2318
2319	2332	2333	2335	2336
2337	2339	2340	2348	2349
2390	2391	2392	2393	2394
2395	2396	2397	2398	2399
V073	V078	V1000	V1001	V1002
V1003	V1004	V1005	V1006	V1007
V1009	V1011	V1012	V1020	V1021
V1022	V1029	V103	V1040	V1041
V1042	V1043	V1044	V1045	V1046
V1047	V1048	V1049	V1050	V1051
V1052	V1053	V1059	V1060	V1061
V1062	V1063	V1069	V1071	V1072
V1079	V1081	V1082	V1083	V1084
V1085	V1086	V1087	V1088	V1089
V109	V580	V581	V661	V662
V671	V672	V711	V760	V7610
V7611	V7612	V7619	V762	V763
V7641	V7642	V7643	V7644	V7645
V7646	V7647	V7649	V7650	V7651
V7652	V7681	V7689	V769	0422

* Claims for prescription drugs were also used to define prevalence of hypertension and diabetes. Codes available upon request.